

1장. H&A본부 CX DX 추진 방향

- 기업별 DX 정의: 디지털 기술 중심의 비즈니스 혁신
- 본부의 DX 정의: 디지털시대, 고객/기술 변화에 대한 이해를 바탕으로 **사업방식을 혁신하는 모든 활동**

[키워드] 패러다임의 변화 / 진화방향

- 소비자 니즈 개인화, 파괴적 혁신(Disruptive Innovation), 경쟁구조 빠른 다변화, 글로벌 금융시장 불안, 코로나19
- 서비스 / 콘텐츠(솔루션), 개인 맞춤형, 온라인

- 데이터: 제품과 서비스가 기획되고, 만들어지고, 판매되고, 사용되면서 발생하는 모든 흔적

고객경험(CEJ): 데이터 기반으로 고객을 이해하기 위한 첫 단계

[3월기술] H&A가 고객 데이터 기반의 제품 기획을 실시하는 이유

CEJ 중요한 이유

- : 경쟁사 접할 기회↑ → 많은 정보량으로 인한 낮은 충성도
- : 기업 경험 기반으로 브랜드 구별
- : 기존 고객 Lock-In으로 비즈니스 성장

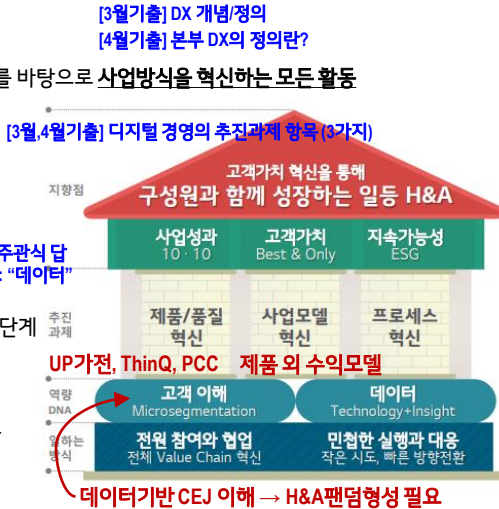
- CX: 모든 고객 접점, 연결, 서비스, 상호작용, Outside-In, Gain Point, 고객의 삶에 대한, 생활의 변화

→ 제품 중심에서 벗어나 **“고객과 상호작용하며 함께 발전”**해나가야 함.

[4월기술] H&A디지털 경영 로드맵: DX Starter(20년) - Literate (21년) - Performer(22년) - Leader(23년) ※오답, Pioneer(22년)

- CX 지향점: Q / C / D → FUN (First / Unique / New) → 고객 감동

일하는 방식의 혁신 고객경험의 혁신 [3,4월기술] 전자CX 지향점의 기본역량 3요소 (정당, FUN 풀네임)



데이터 기반 CEJ 이해 → H&A팬덤형성 필요

2장. H&A 데이터 개요

- 빅데이터: 원천 데이터의 대규모화 / 다양화 / 확장

- 내부데이터: 기기/앱데이터, 고객정보, 상담 및 수리 / 외부데이터: 고객 라이프 스타일, 환경/지역, 고객 VoC

3장. H&A기기/앱로그 데이터

- 기기/앱로그 데이터: ThinQ앱과 기기를 통해 수집되는 데이터

앱로그 데이터 기기데이터

- 기기 데이터: 가전제품에서 발생한 제품의 상태나 동작, 사용자와의 상호작용을 저장하는 데이터

(ex: 고객제어 정보, 고객설정 정보, 환경 정보, 제품상태 정보, 사용이력/누적 정보, 진단 정보)

= 앱 데이터 + 스마트 데이터

★ 앱로그 데이터와 다름!!

- 앱로그 데이터: ThinQ앱에서 발생한 사용자와의 상호 작용을 저장하는 데이터

(ex: Firebase 기본 정보, 등록제품 정보, 화면 정보, 지역 정보, Push정보, 에러 정보 등)

4장. H&A본부 데이터 활용 절차

- 개인정보: LG전자 제품/서비스 이용 흔적 모두, 임직원 정보 포함. 고객정보는 고객의 것 → 사용 동의 필요 [3월기술]

→ 분석시 예외 조건: 통계 / 과학적 / 공익적 목적 → 단, 가명화 필수 [3월, 4월기술]

★ 분석마다 활용 동의를 받아야 하는 건 아님!!

- 유출 피해: 전체 매출액의 3% 부과 [3월기술]

- 데이터 활용 단계: 데이터 요청 > 사전 검토/승인 > 수집 처리 > 검증 > 전달/폐기 > 가공 전 데이터

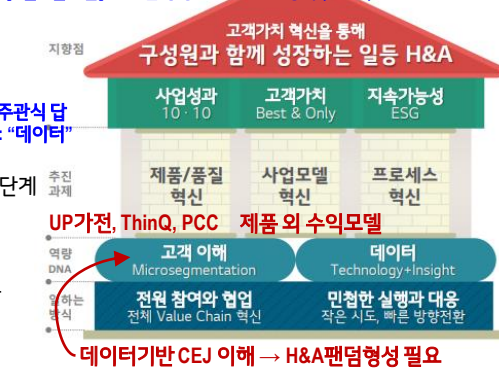
- 데이터 제공 요청 명세서 (@데이터 추출 사전 검토) [4월 기술]

- 추출 데이터 / 수집 기간, 조건 / 목적 / 실사용자 / 사용기간

[3월기술] DX 개념/정의

[4월기술] 본부 DX의 정의란?

[3월, 4월기술] 디지털 경영의 추진과제 항목 (3가지)



데이터 기반 CEJ 이해 → H&A팬덤형성 필요

5장. 데이터 플랫폼의 이해

[3월기술] SSBI + Sandbox 특징

- 1. SSBI (Self-Service BI): 데이터 시각화 → 데이터 분석/활용 용이하게 함.

- 특징: User가 Data에 직접 접근(쉬운 접근성), 직접 분석, 빠른 요구사항 대응
자유로운 데이터 탐색(용량 한도X), 능동적인 데이터 시각화,
뛰어난 데이터 Live Connection (실시간 자동 업데이트)

- 2. Sandbox: 데이터 과학자 및 심층 분석가들이 사용하도록 설계된 분석 플랫폼

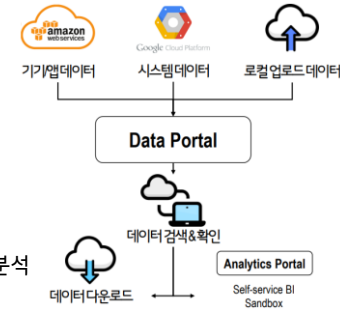
- 특징: 높은 확장성(라이브러리 기반 확장 기능), 어려운 난이도(코딩 지식 필요)

- 사용 방법: 분석 과제 생성 → 분석 환경 생성 → DP의 데이터 조회 → 데이터 분석

- 3. DLMS (Data Lifecycle Management System)

- 목적: 데이터 Life cycle 단계별 데이터 관리, 대시보드 기반 모니터링

- 프로세스: 데이터셋 등록 → 프로젝트 등록 → 데이터셋 맵핑 → 데이터 기획 → 개발 → 검증 → 활용



6장. 데이터의 유형과 기술통계분석

[3월기술] 범주형, 수치형 예시

[3월기술] 범주형(막대/파이), 수치형(히스토그램/상자그림)

[3월기술] 히스토그램-상자그림 대응

[4월기술]

- 통계량: 자료의 특징을 나타내는 요약치

- 분포: 데이터가 가지는 모든 값의 형태를 통칭

- 기술통계분석: 수집한 데이터를 정리, 요약, 해석, 표현하여 자료의 특징을 규명하는 분석방법

- 1. 범주형: 순서가 있는/없는 범주형 [키워드]

- 빈도표 (범주별 데이터 개수 요약), 비율 (각 범주의 빈도수 비율), 막대그래프 / 파이그래프

- 2. 수치형: 이산형 자료 혹은 연속형 자료 [키워드]

★ 평균 하나만으로 자료의 분포를 설명할 수 없음!!

- 평균 (수치형 자료를 요약하는 대표 방법, =Average())

★ 평균과 표준편차는 이상치에 영향을 받지만, 중앙값과 IQR은 영향을 받지 않음!

- 중앙값 (자료의 중심 대푯값, =Median())

- 최빈수 (데이터에 가장 자주 나타나는 값, =Mode())

[4월기술] 엑셀로 푸는 문제 3개 (DD_PDT_100기준)

1) 토요일에 에어컨 사용한 날 (정답: 4. 주관식)

2) 온도별 정수기 냉수취수량 그래프 (정답: 산점도 형상 찾기, 객관식)

3) 정수취수량의 중앙값 (정답: 1594)

[4월기술] 객체의 다양한 유형, 빈도를 담은 집합을 분석하는 틀 (히스토그램)

- 분산 (자료의 퍼짐 대푯값)

- 표준편차 (분산의 제곱근 값 =STDEV.S())

- 범위 (최대값 빼기 최소값, =MAX()-MIN())

- IQR (사분위수 범위, =QUARTILE.EXC.(3)-QUARTILE.EXC.(1))

★ Upper whisker, Q3, 중앙값, Q1, Lower whisker [3월기술]

- 히스토그램 (빈도를 나타낸 그래프), 상자그림 (다섯수치요약과 이상치를 시각화)

- 데이터의 치우침: 긴 오른쪽(중앙값<평균), 긴 왼쪽(중앙값>평균), 대칭분포 (중앙값=평균)

- 3. 수치형 vs. 수치형: 수치형 두 변수의 상관성을 시각화 → 산점도, 상관관계 (상관관계: 인과성 설명 불가, 상관성 Only)

※ 이상치로 상관계수가 크게 나타날 수 있음 → 산점도와 상관계수 함께 해석 필요

- 4. 범주형 vs. 수치형: 그룹에 따른 수치형 자료 비교/분석 목적

※ 평균 차이가 큰 두 그룹이라도 같은 분포 집단일 수 있음 → 평균/표준편차 + 상자그림 or 히스토그램 확인 필요

- 5. 범주형 vs. 범주형: 두 범주형 자료 요약 → 분할표 / 막대그래프 / 모자이크 그림

- 6. 시간에 따른 수치형 자료의 시각화: 경향성 확인

[4월기술] 세탁기의 포량과 운전시간에는 어떤 관계가 있을까?

(정답: 수치형 - 수치형)

[4월기술] 히스토그램, 상자그림 해석 문제 (산포, 평균 등등)

[4월기술] 데이터 분석 과정 순서대로 바르게 나열한 것은?

분석과정과 각 과정에 대한 설명이 틀리게 짝지어진 것은?

7장. H&A본부 데이터 분석 방법론 소개

- 데이터 분석 과정

- 문제 정의: 현실의 구체적인 문제를 명확하게 표현 → 통계적/수리적 언어로 번역

- 분석 기획: 변수/지표 등을 정의

[3월기술] 데이터 분석 과정 중 '문제 정의' 단계

- 데이터 취득: 다양한 형태/시스템의 원데이터를 분석 시스템으로 가져오는 작업 (개인정보 Compliance 검토)

- 데이터 가공: 분석표로 가공

- 데이터 분석: EDA, Modeling [3월, 4월기술] EDA (탐색적 데이터 분석): 시각화와 간단한 통계량을 통해 데이터 패턴 발견, 이상치 발견

15. 다음 중 LG Data Portal에서 기기데이터를 다운로드 받을 시 옳은 것은? [매월 5점]
데이터항목의 특정값으로 필터링하여 다운로드 받을 수 있다.
한번에 일련성 데이터들 다운로드 받을 수 있다.
데이터셋에서 쿼리(데이터베이스에 정보를 요청)를 하여만 다운로드 받을 수 있다.
데이터셋의 필드는 항목을 선택하여 다운로드 받을 수 없다.